STUDY OF MOMENTS ON PARETO-II DISTRIBUTION

BY

ODION AUGUSTINE SYLVESTRE

MAT NO: PSC1607802

BEING A PROJECT PRESENTED TO THE DEPARTMENT OF STATISTICS, FACULTY OF PHYSICAL SCIENCES, UNIVERSITY OF BENIN, BENIN CITY IN PARTIAL FUFILMENT OF THE REQUIREMENTS OF THE AWARD OF BACHELOR OF SCIENCE (B.SC) IN STATISTICS.

JULY 2021

CERTIFICATION

We certify that this project work was carried out by ODION AUGUSTINE SYLVESTRE with MAT NO: PSC1607802 of the Department of Statistics, Faculty of Physical Sciences, University of Benin, Benin City. We examined and found it acceptable for the award of Bachelor's Degree in Statistics, University of Benin.

______________________________ ______________________

DR. PATRICK OSATOHANMWEN **DATE**
Project Supervisor

______________________________ ______________________

PROF. C.C ISHIEKWENE **DATE**
Head of Department

______________________________ ______________________

External Examiner **DATE**

1/9/2021

UNDERTAKING

This project work was carried out by me ODION AUGUSTINE SYLVESTRE and has not copied the work of any other author. All works used have duly cited and acknowledged.

ODION AUGUSTINE SYLVESTRE

DATE

DEDICATION

This research project is dedicated first to ALMIGHTY GOD for his enabling strength he gave me in completing this work. Secondly to my Late father MR. DICKSON ODION and lastly to my Guardians/Sponsors MR. INNOCENT ODION & MR. OHIS ODION for their support financially and otherwise throughout the course of this study.

ACKNOWLEDGEMENT

I am grateful to God almighty, the provider of wisdom and knowledge for his love, mercies and grace throughout the period of my programme.

I sincerely appreciate my supervisor DR. PATRICK OSATOHANMWEN for availing his time in actualization of this project, for his timely criticism and corrections that led me through the various stages of this project.

I also appreciate my brothers who are also my sponsors MR. ODION INNOCENT and MR. OHIS ODION for their advice and support financially, mentally and otherwise from day one till this day.

Lastly, not forgetting my friends and coursemates who have contributed one way or the other towards my stay in this Institution, EMEKUME CHARLES OLISEMEKE for his support academically, OSARONOMA ENDURANCE, OSAYANDE ELIZABETH and a host of them for their unquantifiable love and kindness.

May the good Lord bless you all in Jesus name. Amen

TABLE OF CONTENTS

ABSTRACT

This work is intended to study the moments of distributions and in particular Pareto-II distribution named after Italian Scientist, Vilfredo Pareto. This study was guided by the following objectives; to obtain the rth moment, obtain the mean, variance, skewness and kurtosis of the Pareto-II distribution and finally obtain some numerical results of the moment.

The study employed the knowledge of differential and integral calculus with transformation of variables to obtain several expressions as we shall be seeing. Statistical software "R" was used to run analysis and obtain numerical results of the moments.

Findings revealed that the parameters of the distribution are important in determining the behavior of the moments. From the findings, it implied that the study of moments is important and applicable to study of distributions.

Keywords: Moments, Parameter, Distribution, Mean, variance, Skewness, kurtosis.

CHAPTER ONE

INTRODUCTION

1.0. BACKGROUND OF THE STUDY

Like every other works of Statistics, the knowledge and understanding of moments cannot be underestimated as it finds it application in Random Variables. This section of this chapter highlights an introduction to moments of random variables with respect to Pareto Type II distribution.

In mathematics, the moments of a function are quantitative measures related to the shape of the functions' graph. The concept is used in both Mechanics and Statistics. The Moment of random variables (or of it distribution) are expected values of powers or related function of the random variable. The moment of a random variable can be computed by using either its Moment Generating Function, if it exist or it characteristics function.

In studying moments of random variable, we define the "r^{th}" moment. The "r^{th}" moment of a random variable say "x" about the origin is defined by

$$E(X^r) = \begin{cases} \sum_{x=0}^{n} X^r\, P(x) & \text{when x is discrete} \\ \int_{-\infty}^{\infty} X^r f(x)dx & \text{when x is continuous} \end{cases}$$

In particular, if the first moment i.e r=1 is taken, then this becomes the mean denoted as $\mu_x = E(x)$ which is the center of location of a distribution.

Also, if r=2, then this becomes the second moment i.e $E(X^{2)}$ which is known as the variance. If r=3, i.e $E(X^3)$ we refer to the third central moment which is the measure of the lopsidedness of the distribution, any symmetric distribution will have a third central moment, if defined, of Zero. A distribution that is skewed to the left will have a negative skewness while a distribution that is skewed to the right will have a positive skewness. If the fourth central moment is taken, i.e r=4 is a measure of heaviness of the tail of the distribution

In mathematical Statistics, there are different types of distribution but in the course of this study, our focus is to study the moment of a particular distribution known as the Perato type II Distribution.

Pareto distributions are the most popular models in economics, finance and related areas. In fact, the first Pareto distribution due to Pareto was used to model the allocation of wealth among individuals. Since Pareto, several extended Pareto distributions have been proposed in the literature and have been applied in a wide variety of fields. The list of applications is too exhaustive, however, some recent applications have included: income modeling the wealth distribution in the Forbes 400 list, commercial fire loss severity in Taiwan, city size distribution in the United States. Pareto distributions are increasingly being used to model problems in economics and finance. Hence, it is essential to have tools to check the goodness of fit (GOF) of Pareto distributions. Several tests have in fact been proposed to check the GOF of Pareto distributions, however, we are not aware of any review covering all known tests for Pareto distributions. Such a review is essential for practitioners given the wide spread use of Pareto distributions and such a review could also encourage the development of more GOF tests.

The Perato-II also known as Lomax distribution was named after Italian Economist and sociologist Vilfredo Pareto. It is sometimes referred to as the Pareto Principle or "80-20" rule. The Pareto Distribution was first employed in Italy in the 20th Century to describe the distribution of wealth among the population, it is used in describing Social, Scientific and geographical phenomena in society. It definition was expanded in the 1940s by Dr. Joseph M. Juren. In the 1890s, Vilfredo Pareto studied income tax data from several countries. He plotted the number of people earning an income above a certain threshold against the respective threshold on double logarithmic paper and revealed a linear relationship. The corresponding income distribution was more skewed and heavy-tailed than bell-shaped curves: Pareto felt that he had discovered a new type of "universal law" that was the result of underlying economic mechanisms. Since then, Pareto's discovery has been confirmed and generalized to the distribution of firm size (Axtell (2001) and wealth, which also follow Pareto distributions in the upper tail. According to Gabaix (2016), this is one of the few quantitative "laws" in economics that hold across time and countries. A striking qualitative feature of Pareto tails is large inequality: for example, the top 1% gets about 20% of pre-tax income in the United States. Therefore, understanding the source of Pareto distributions is a first order question for the economics of inequality

1.1. AIM AND OBJECTIVES

The aim of this study is to examine the moments of the Pareto-II Distribution

The objectives are:

- **i.** To obtain the r^{th} moment
- **ii.** Obtain the Mean, Variance, Skewness, and Kurtosis
- **iii.** To obtain some empirical results of the moment

1.2. SIGNIFICANCE OF THE STUDY

The importance of the study of moments to distributions cannot be overemphasized as it finds many applications and uses in the study of distributions. Moments are very useful in Statistics because they tell you more about your data and obtain some certain information. There are four common applications of moments to distributions viz: Mean, variance, skewness and kurtosis. The mean gives a measure of centre of the distribution when r=1, the square root of the variance is the standard deviation, and this explains the dispersion of the data about the mean when r=2. Skewness describes the shape of the distribution for r=3, and kurtosis measures the peakedness or flatness of the distribution if r=4

1.3. DEFINITION OF TERMS

Random Variable: In probability and statistics, a random variable or stochastic variable is described informally as a variable whose values depend on outcomes of a random phenomenon. It is a variable whose value is unknown or function that assigns values to each experiment's outcomes

Distribution: A distribution in Statistics is a function that shows the possible values of a variable and how often they occur.

Moments: Moments are set of statistical parameters to measure a distribution.

1.4. STRUCTURE OF THE STUDY

This section emphasizes on the structure of this study, it gives a summary of what the reader expects as they progress through the pages of this work. Chapter One gives an introduction to the study, discussing the moments of random variable and an introduction to the Pareto-II distribution. It also highlights the aim and objectives of the study, the significance of the study and definition of some basic terms.

The next chapter (Chapter Two) will be reviewing literature that are relevant to the study of Perato-II Distribution. In Chapter we shall take a look at methodologies of the study, chapter IV we look at Data Analysis and Application.

The Study is concluded in chapter V with a summary of all the preceding chapters and also a general conclusion of the study.

CHAPTER TWO

LITERATURE REVIEW

2.0 INTRODUCTION

The Pareto Distribution named after Italian Sociologist and Economist Vilfredo Pareto was first employed in Italy in the early 20th Century to describe the distribution of wealth among the population. Therefore it is pertinent to do a review of literature on this study and it applications

2.1 LITERATURE REVIEW ON PARETO DISTRIBUTION

According to Pareto (1895), social institutions could not be the underlying reason for these regularities, as they were observed in very different societies. He also dismissed random chance, as chance does not produce such thick tails. He concluded that Pareto distributions must arise from "human nature." The modern economics literature has used either random growth models or the distribution of primitives to explain the emergence of Pareto distributions. In random growth models, the stochastic process is assumed to be scale independent (Gibrat's (1931) law), and one looks for stationary distributions created by that process. Gibrat's law also intuitively leads to scale independence in the stationary distribution created by the process – thus a power law distribution – and Zipf's (1949) law when frictions become small.

Vilfredo Pareto's *magnum opus* was his (1916) 'Treatise on General Sociology'. This was written, for the most part, before the First World War, although it was not published in English until 1935, twelve years after his death. Writing in 1966, the political scientist Sammy Finer rated this as 'the most pregnant work of political science in the last half century', because be believed it provided a rich source of testable hypotheses for future researchers (Pareto 1966, 87). This was Pareto's intention. He felt that it made good sense, for scientists and historians alike, to theorise with generalities before moving to specifics (Pareto 1935, §144, §540). Hence 'general sociology' had to precede the various 'special sociologies' such as political sociology. Pareto believed that once he had established a general framework theory for sociological investigation, social scientists following after him could begin to contribute detail to his theory at micro sociological levels more amenable to empirical enquiry. He valued his own general theory with some humility, hoping others would rework its constituent parts to accommodate future research findings

In 1974 Salem and Mount suggested that a simple gamma distribution might be adequate to fit income distributions. Cramer (1978) pointed out that it fares quite well when compared to the log-logistic (i.e., Pareto (III) with $\mu = 0$). Singh and Maddala (1976) had proposed the Pareto (IV) model using an argument involving decreasing failure rates

Singh and Maddala (1975, 1978) proposed a rich family of Lorenz curves for fitting purposes. This family included the Weibull, logistic and Pareto IV curves. These suggestions further augmented the list of possible distributions one might consider in fitting income data

Kloek and Van Dijk (1978) provided an attempt to empirically select a suitable distribution from the wide variety available. It is interesting that they still failed to fit the data well. Perhaps we must fall back to Mandelbrot's observation that, since for low incomes the typical distribution is erratic, it is unlikely that a single theory can account for all features of income distribution. Rather than throw up our hands in dismay, we may well go right back to Pareto and restrict efforts to modeling only the upper tail of the income distribution

The Lomax distribution is known as a special form of Pareto Type - II distribution (Pareto, 1898), the cumulative density function of Pareto Type - II is given by:

$$F(x) = \left[1 + \frac{x-\mu}{\sigma}\right]^{-\alpha}, x \geq \mu$$

Specifically the Lomax distribution is a Pareto type II distribution with $X_m = \lambda$ and $\mu = 0$, It is also a mixture of exponential and gamma distributions. It belongs to the family of decreasing failure rate in the lifetime context, (Chahkandi and Ganjali 2009) and also arises as a limiting distribution of residual lifetimes at great age (Balkema and De Haan 1974). This distribution has been suggested as a heavy-tailed alternative to the exponential, weibull and gamma distributions. Furthermore, It is related to the Burr family of distributions (Tadikamalla, 1980) and as a special case, can be obtained from compound gamma distributions (Dubey, 1970). Harris (1968) used Lomax distribution to analyze income and wealth data. Atkinson and Harrison (1978) also used it for modeling business data, while Corbelini et al (2007) used it to model firm size and queuing problems. Lomax distribution is used as the basis for several generalizations, for example Al-Awadhi and Ghitany (2001) used Lomax distribution as a mixing distribution for the Poisson parameter and derive a discrete Poisson-Lomax distribution and Punathumparambath (2011) introduced the double-Lomax distribution and applied it to IQ data.

Cordiero et al (2014) proposed a new generalized Lomax family of distributions called the generalized Lomax – G family with two extra positive parameters to generalize any continuous baseline distributions. Some special model such as Lomax – Normal, Lomax – Weibull, Lomax – Logistic and Lomax – Pareto distributions were discussed. Some mathematical properties of the new generalization including ordinary and incomplete moments. Quantile and generating functions, mean and mean deviations distribution of order statistics were presented. They discuss the estimation of model parameters using the Method of Maximum Likelihood. They also defined a Log – Lomax – Weibull regression model for censored data.

Osagie and Nosakhare (2020) introduced a new three parameter Lomax-Exponential distribution. They presented its defining functions and statistical properties such as quantile function, moments, inequality curves, entropy measures, measures of residual life and order statistics were discussed. Inferences for point and interval estimation for the parameters of the Lomax – Exponential distribution were presented. Application of the new distribution to lifetime data is illustrated to determine the usefulness and applicability in lifetime analysis.

Champernowne (1953) is perhaps the first of such random growth models for incomes, and Simon and Bonini's (1958) for firms. Kesten (1973), Gabaix (1999), and Luttmer (2007) are other examples of this approach. There is also a literature that links hierarchies to power law distributions, in the case of both firms and cities – for example Lydall (1959) for firms and Beckmann (1958) for cities. Hsu (2012) is a micro foundation of Beckmann (1958) using central place theory, and in which a multiplicative process occurs at the spatial level rather than in a time dimension. Another way the literature has generated Pareto distributions is by using Pareto as primitives' distribution, such as Lucas (1978), Helpman et al. (2004), Chaney (2008), Terviö (2008) and Gabaix and Landier (2008). The boundaries of the firm are defined as in the span of control model of Lucas (1978), who first formalizes that the limits to the boundaries of the firm can arise from limited managerial attention. Garicano (2000) is more explicit about what management is; his model leads to the kind of production functions that I emphasize in this paper.4 Garicano and Rossi-Hansberg (2006) is certainly the most closely related paper: They investigate the implications of the Garicano (2000) model on the distribution of incomes, but make no mention of Pareto distributions.

Pickands (1975) introduced the generalized Pareto (GP) distribution and has since been applied to a number of areas including socio-economic phenomena, physical and biological processes (Saksena & Johnson, 1984), reliability studies and the analysis of environmental extremes. Davison & Smith (1990) pointed out that the GP distribution might form the basis of a broad modelling approach to high-level exceedances. DuMouchel (1983) applied it to estimate the stable index a to measure tail thickness, whereas

Davison (1984a, 1984b) modelled contamination due to long-range atmospheric transport of radionuclides, van Montfort & Witter (1985, 1986) and van Montfort & Otten (1991) applied the GP distribution to model the peaks over a threshold (POT) stream flows and rainfall series, and Smith (1984, 1987, 1991) applied it to analyse flood frequencies and wave heights. Similarly, Joe (1987) employed it to estimate quantiles of the maximum of iv observations. Wang (1991) applied it to develop a POT model for flood peaks with Poisson arrival time, whereas Rosbjerg et al. (1992) compared the use of the 2-parameter GP and exponential distributions as distribution models for exceedances with the parent distribution being a generalized GP distribution. In an extreme value analysis of the flow of Burbage Brook, Barrett (1992) used the GP distribution to model the POT flood series with Poisson inter-arrival times. Davison & Smith (1990) presented a comprehensive analysis of the extremes of data by use of the GP distribution for modelling the sizes and occurrences of exceedances over high thresholds. Methods for estimating the parameters of the 2-parameter GP distribution were reviewed by Hosking & Wallis (1987).

Quandt (1966) used the method of moments (MOM), while Baxter et al (1980, 1981) used the method of maximum likelihood estimation (MLE) for the Pareto distribution. The MOM, MLE and probability weighted moments (PWM) were included in the review, van Montfort & Witter (1986) used the MLE to fit the GP distribution to represent the Dutch POT rainfall series and used an empirical correction formula to reduce bias of the scale and shape parameter estimates. Davison & Smith (1990) used the MLE, PWM, a graphical method and least squares to estimate the GP distribution parameters. Wang (1991) derived the PWM for both known and unknown thresholds.

Kotz and Balakrishnan (1994) and Rodriguez (1977). In most of these references the distributions are labeled Burr distributions. The moments of the classical Pareto distribution were known in the nineteenth century. It is only in the Pareto (I) and (II) cases that simple expressions are obtainable for the variance. The unpleasant form of the expressions for the variances of the various distributions should not be cause for alarm. In the context of modeling income distributions, the variance is not often of importance. It is much more important that other measures of inequality be related in as simple a manner as possible to the parameters in the distribution

Tadikamalla (1980) provides discussion and a diagram displaying the range of possible values of the skewness and kurtosis measures for the Pareto (IV) distribution.

Seal (1980) provides an expression for involving the exponential integral function.

Nadarajah and Kotz (2006) give an expression in terms of the Whittaker function, while Takano (2007) presents an expression involving confluent hypergeometric functions

Al et al (2003) presented the geometrical properties of the Pareto Distribution

Ismail (2004) discussed a simple estimator of the shape parameter of the Pareto distribution.

Thorin (1977) showed that the Pareto (II) $(0,\sigma,\alpha)$ distribution is a generalized transform and thus is infinitely divisible. A shorter alternative proof was provided by Berg (1981).

Blum (1970) considered the case of a sum of n independent Pareto I $(\alpha,1)$ random variables. He showed that an asymptotic expansion obtained. Blum devoted considerable attention to the form of these coefficients. The expression is probably only of technical interest. It, obviously, has potential as a tool for identifying the limiting distribution for sums of Pareto variables. Blum made use of it in that context.

Ramsay (2003) provides an alternative expression for fn(x) in terms of exponential integral functions for the case in which α is a positive integer

Charles Powers (1987) reaffirms Finer's estimation that it is Pareto's broad sociological framework theory which retains most value. He argues that this has a timeless relevance because it 'provides lessons about the social structural dynamics which have operated throughout human history' (Powers 1987, 11). As will shortly be explained, Powers tries to distil his theory and restate its key claims as a detailed set of empirical propositions. In doing this, he felt he was completing Pareto's sociological project (Powers 1987, 12). Powers' restatement of Pareto argues that 'social sentiment', 'economic organization' and 'political organization' are each characterized by cyclical change. His 'elementary theory' for each of these cyclical processes consists of a set of interlocking mechanisms. Each of these mechanisms has a similar structure, whereby it is claimed that one kind of change is likely to induce another kind of change (if Δ A then Δ B). Then Powers lists further mechanisms which explain how the endogenous dynamics of the social, political and economic cycles are likely to impact exogenously upon each other. In other words, we are given detailed explanations for how economic change is likely to influence the direction of social and political change, how social change is likely to influence the direction of political and economic change, and how political change is likely to influence the direction of social and economic change

Gupta et al. extended the Pareto distribution by raising $G(x{:}\beta,k) = 1 - \left(\frac{\beta}{x}\right)^k$ to a positive power. In this note, we refer to this extension as the exponentiated Pareto (EP) distribution. Recently, many authors have considered various exponentiated-type distributions based on some known distributions such as the exponential, Rayleigh, Weibull, gamma and Burr distributions; see, for example, Gupta and Kundu and, Surles and Padgett, Kundu and Raqab and Silva et al. The methods of moments and maximum likelihood have been used to fit these models

Akinsete et al and Mahmoudi extended the Pareto and GP distributions by defining the beta Pareto (BP) and beta generalized Pareto (BGP) distributions, respectively, based on the class of generalized (so-called "beta-G") distributions introduced by Eugene et al. The generalized distributions are obtained by taking any parent G distribution in the cdf of a beta distribution with two additional shape parameters, whose role is to introduce skewness and to vary tail weight. Following the same idea, many beta-type distributions were introduced and studied, see, for example, Barreto-Souza et al, Silva et al and Cordeiro et al.

2.2. SUMMARY

In this chapter, we have been able to review some relevant literatures that are essential to this study. It is obvious that a lot of works have been carried out on the topic and it applications to different field of study

CHAPTER THREE

METHODOLOGY

3.0. INTRODUCTION

In this chapter, we want to explore the moment of the Pareto Type II-distribution. The aim is to find a functional form for the moment, obtain the first, second, third and fourth moment.

3.1 PARAMETER ESTIMATION OF PARETO-II DISTRIBUTION

Maximum Likelihood Estimation Method

As observed by the properties and the characteristics of the Pareto-II distribution, the parameters; shape and scale parameter have certain behavior which can be estimated.

Once a model is specified with its parameter and data collected, one is in a position to evaluate its goodness-of-fit that is if it fits the observed data.

The Maximum Likelihood Estimation Method (MLE) is used to estimate the parameters of the Pareto-II distribution.

The Pareto-II probability density function (PDF) is given as

$$f(x;\alpha,\lambda) = \begin{cases} \frac{\alpha}{\lambda}\left(1+\frac{x}{\lambda}\right)^{-(\alpha+1)} \\ 0 \quad elswehere \end{cases}$$

Applying the Maximum Likelihood Estimation to estimate the two – parameters (shape α and scale λ parameter).

Let $x_1, x_{2,\dots} x_n$ be a random sample of Pareto-II distribution, Then the Likelihood function is expressed as:

$$L(x_1, x_{2,\dots} x_i; \alpha, \lambda) = \prod_{i=1}^{n} \left(\frac{\alpha}{\lambda} \left(1 + \frac{x_i}{\lambda}\right)^{-(\alpha+1)} \right)$$

Where $i = 1(1)n$

$$L(x_i;\ \alpha, \lambda) = \left(\frac{\alpha^n}{\lambda^n}\right) \sum_{i=0}^{n} \left(1 + \frac{x_i}{\lambda}\right)^{-(\alpha+1)}$$

Taking the Log - Likelihood function, to obtain;

$$LogL = nlog\alpha - nlog\lambda - (\alpha + 1) \sum_{i=1}^{n} log\left[1 + \frac{x_i}{\lambda}\right]$$

Differentiating $LogL$ with respect to α, λ that is $\frac{\partial LogL}{\partial \alpha}$, $\frac{\partial LogL}{\partial \lambda}$ and equate to zero.

$$\frac{\partial LogL}{\partial \alpha} = \frac{n}{\alpha} - \sum_{i=0}^{n} \log\left[1 + \frac{x_i}{\lambda}\right]$$

$$\frac{n}{\breve{\alpha}} - \sum_{i=0}^{n} \log\left[1 + \frac{x_i}{\lambda}\right] = 0$$

$$\frac{n}{\breve{\alpha}} = \sum_{i=0}^{n} \log\left[1 + \frac{x_i}{\lambda}\right]$$

$$\breve{\alpha} = \frac{n}{\sum_{i=0}^{n} \log\left[1 + \frac{x_i}{\lambda}\right]}$$

$$\frac{\partial LogL}{\partial \lambda} = -\frac{1}{\lambda} + \frac{(\alpha + 1)}{n\lambda} \sum_{i=0}^{n} \left[\frac{x_i}{\lambda + x_i}\right]$$

3.2 THE CUMMULATIVE DENSITY AND PROBABILITY DENSITY FUNCTIONS

We shall obtain the Probability Density Function of the Pareto Type II distribution given it cumulative Density Function

It Cumulative Density Function (CDF) is given as

$$F(x) = 1 - \left(1 + \frac{x}{\lambda}\right)^{-\alpha}$$

And it Probability Density Function (PDF) is given as

$$f(x) = \frac{\alpha}{\lambda}\left(1 + \frac{x}{\lambda}\right)^{-(\alpha+1)}$$

Proof

To obtain a Probability Density Function, we differentiate it Cumulative Density Function,

$$\frac{d}{dx}\left[1 - \left(1 + \frac{x}{\lambda}\right)^{-\alpha}\right]$$

Using Chain's rule $\frac{dy}{dx} = \frac{dy}{du} \times \frac{du}{dx}$ to differentiate

Let $u = 1 + \frac{x}{\lambda}$

$$\frac{du}{dx} = \frac{1}{\lambda} \qquad *$$

Recall that

$$y = 1 - \left[1 + \frac{x}{\lambda}\right]^{-\alpha}$$

$$=> y = 1 - u^{-\alpha}$$

$$\frac{dy}{du} = -(-\alpha)u^{-\alpha-1}$$

$$= \alpha u^{-\alpha-1} \qquad **$$

Inserting * and ** into the chain rule above we get

$$\frac{dy}{dx} = \frac{dy}{du} \times \frac{du}{dx} = \alpha u^{-\alpha-1} \times \frac{1}{\lambda}$$

$$= \frac{1}{\lambda}\alpha u^{-\alpha-1}$$

$$= \frac{\alpha}{\lambda}u^{-\alpha-1}$$

Recall that $u = 1 + \frac{x}{\lambda}$

$$\therefore f(x) = \frac{\alpha}{\lambda}\left[1 + \frac{x}{\lambda}\right]^{-(\alpha+1)} \qquad ***$$

Equation (***) is the Probability Distribution Function (PDF) of the Pareto type II distribution (Q.E.D)

3.3 THE RAW MOMENT

We wish to obtain the raw moment of the Pareto type II distribution. The raw moment of any given continuous distribution is given as

$$U_r^1 = E(X^r) = \int_0^\infty f(x)dx$$

$$U_r^1 = \int_0^\infty x^r \left[\frac{\alpha}{\lambda}\left[1+\frac{x}{\lambda}\right]^{-(\alpha+1)}\right] dx$$

$$U_r^1 = \frac{\alpha}{\lambda}\int_0^\infty x^r \left[1+\frac{x}{\lambda}\right]^{-(\alpha+1)} dx$$

$$U_r^1 = \frac{\propto}{\lambda}\int_0^\infty \frac{x^r}{\left[1+\frac{x}{\lambda}\right]^{(\propto+1)}} dx \qquad 1$$

Recall the Beta Function given as

$$\beta(x,y) = \int_0^\infty \frac{x^{y-1}}{[1+t]^{x+y}} dt \qquad 2$$

Comparing Equation 1 and 2 we have that

$$t = \frac{x}{\lambda} \quad => x = \lambda t$$

$$\therefore \quad \frac{dx}{dt} = \lambda \quad => \quad dx = \lambda dt$$

Applying the change of variables above to make equation 2 look like 1. From equation 1 above becomes

$$U_r^1 = \frac{\alpha}{\lambda}\int_0^\infty \frac{(\lambda t)^r}{(1+t)^{(\propto+1)}} \times \lambda \; dt$$

$$U_r^1 = \frac{\alpha}{\lambda}\int_0^\infty \frac{\lambda^r t^r}{(1+t)^{(\alpha+1)}} \times \lambda \, dt$$

Simplifying further

$$U_r^1 = \frac{\alpha\lambda^r \lambda}{\lambda}\int_0^\infty \frac{t^r}{(1+t)^{(\alpha+1)}} dt$$

$$U_r^1 = \alpha\lambda^r \int_0^\infty \frac{t^r}{(1+t)^{(\alpha+1)}} dt \qquad 3$$

Equation 3 is the intended equation. Comparing equation 2 and 3 observe that

$$r = y - 1 \Rightarrow y = r + 1$$

Also $(\alpha + 1) = (x + y)$

$$\Rightarrow (\alpha + 1) = (x + r + 1)$$

$$\Rightarrow \alpha + 1 = x + r + 1 \qquad ****$$

Making x the subject of the equation in **** becomes

$$x = \alpha + 1 - r - 1$$

$$x = \alpha - r$$

$$\therefore x = \alpha - r \text{ and } y = r + 1$$

Recall the Beta function stated above in equation 2

$$\beta(x,y) = \int_0^\infty \frac{x^{y-1}}{[1+t]^{x+y}} \quad dt$$

$$\beta(x,y) = \beta(\alpha - r, \quad r+1) = \int_0^\infty \frac{t^r}{[1+t]^{(\alpha+1)}} dt$$

$$U_r^1 = \alpha\lambda^r \int_0^\infty \frac{t^r}{[1+t]^{(\alpha+1)}} dt$$

$$U_r^1 = \alpha\lambda^r \beta\big((\alpha - r), r+1\big) \qquad 4$$

Equation 4 is the raw moment of the Pareto type II distribution

Equation 4 can be further expressed as a relationship between the Beta and Gamma function i.e

$$\beta(x,y) = \frac{\Gamma x \ \Gamma y}{\Gamma x + y}$$

$$\beta\big((\alpha - r), r + 1\big) = \frac{\Gamma(\alpha - r)\ \Gamma r + 1}{\Gamma(\alpha - r) + (r + 1)}$$

$$\therefore \ U_r^1 = \alpha\lambda^r \left[\frac{\Gamma(\alpha - r)\ \Gamma r + 1}{\Gamma\alpha - r + r + 1}\right]$$

$$U_r^1 = \alpha\lambda^r \left[\frac{\Gamma(\alpha - r)\ \Gamma r + 1}{\Gamma(\alpha + 1)}\right] \qquad 5$$

3.4. THE FIRST, SECOND, THIRD AND FOURTH MOMENTS

We can obtain the properties of the distribtion i.e the first, second, third and fourth moments otherwise known as the mean, variance, skewness and kurtosis repesctively from equation 5 above.

The First Moment: r=1 (Mean)

To obtain the mean, we simply put $r = 1$ in equation 5

From equation 5 above, if $r = 1$ becomes

$$U_1^1 = \alpha\lambda \left[\frac{\Gamma(\alpha - 1)\Gamma 1 + 1}{\Gamma(\alpha + 1)}\right]$$

$$= \alpha\lambda \left[\frac{\Gamma(\alpha - 1)\Gamma 2}{\Gamma(\alpha + 1)}\right]$$

$$\Gamma 2 = 1\Gamma 1 = 1$$

$$\Gamma(\alpha - 1) = (\alpha - 1 - 1)! = (\alpha - 2)!$$

$$\Gamma(\alpha + 1) = (\alpha + 1 - 1)! = \alpha!$$

$$U_1^1 = \alpha\lambda\left[\frac{(\alpha-2)!}{\alpha!}\right] = \alpha\lambda\left[\frac{(\alpha-2)}{\alpha(\alpha-1)(\alpha-2)}\right]$$

$$\therefore U_1^1 = \frac{\lambda}{\alpha - 1}$$

This is the first moment or mean of the Pareto Type II distribution

<u>The Second Moment: r=2 (Variance)</u>

Similarly, we put *r=2* in equation 5 to obtain the variance of the distribution

$$U_2^1 = \alpha\lambda^2\left[\frac{\Gamma(\alpha-2)\Gamma 2+1}{\Gamma(\alpha+1)}\right]$$

$$U_2^1 = \alpha\lambda^2\left[\frac{\Gamma(\alpha-2)\Gamma 3}{\Gamma(\alpha+1)}\right]$$

$$\Gamma 3 = 2$$

$$\Gamma(\alpha - 2) = (\alpha - 2 - 1)! = (\alpha - 3)!$$

$$\Gamma(\alpha + 1) = (\alpha + 1 - 1)! = \alpha!$$

$$U_2^1 = \alpha\lambda^2\frac{(\alpha-3)!\times 2}{\alpha!}$$

$$U_2^1 = 2\alpha\lambda^2\frac{(\alpha-3)}{\alpha(\alpha-1)(\alpha-2)(\alpha-3)}$$

Dividing like terms so that

$$U_2^1 = \frac{2\lambda^2}{(\alpha-1)(\alpha-2)}$$

But variance = $E(X^2) - [E(X)]^2$

$$U_2^1 - [U_1^1]^2$$

Where is the Mean found earlier $U_1^1 = \dfrac{\lambda}{\alpha - 1}$

$$[U_1^1]^2 = \left(\frac{\lambda}{\alpha - 1}\right)^2 = \frac{\lambda^2}{(\alpha - 1)^2}$$

$\therefore$ variance = $U_2^1 - [U_1^1]^2$

$$\frac{2\lambda^2}{(\alpha - 1)(\alpha - 2)} - \frac{\lambda^2}{(\alpha - 1)^2}$$

$$\frac{\lambda^2}{(\alpha - 1)}\left(\frac{2}{(\alpha - 2)} - \frac{1}{(\alpha - 1)}\right)$$

$$\frac{\lambda^2}{(\alpha - 1)}\left(\frac{2(\alpha - 1) - (\alpha - 2)}{(\alpha - 2)(\alpha - 1)}\right)$$

Simplifying to get

$$\frac{\lambda^2}{(\alpha - 1)} \times \frac{\alpha}{(\alpha - 2)(\alpha - 1)}$$

$$\frac{\alpha\lambda^2}{(\alpha - 1)^2(\alpha - 2)} \qquad \alpha > 2$$

This is the the Second moment of the Pareto Type *II distribution* otherwise knownn as the variance

The Third Moment $r = 3$ (Skewness)

Similarly to obtain the third moment, we set $r = 3$_ in equation 5 above and simplify.

Now inserting r=2 in Equation 5

$$U_3^1 = \alpha\lambda^3 \left[\frac{\Gamma(\alpha-3)\Gamma3+1}{\Gamma\alpha+1}\right]$$

$$U_3^1 = \alpha\lambda^3 \left[\frac{\Gamma(\alpha-3)\Gamma4}{\Gamma\alpha+1}\right]$$

$$\Gamma4 = 3\Gamma3 = 6$$

$$\Gamma(\alpha-3) = (\alpha-3-1)! = (\alpha-4)!$$

$$\Gamma(\alpha+1) = (\alpha+1-1)! = \alpha!$$

$$U_3^1 = 6\alpha\lambda^3 \ \frac{(\alpha-4)!}{\alpha!}$$

$$U_3^1 = 6\alpha\lambda^3 \ \frac{(\alpha-4)}{\alpha(\alpha-1)(\alpha-2)(\alpha-3)(\alpha-4)}$$

Dividing to take out like terms

$$U_3^1 = \frac{6\lambda^3}{(\alpha-1)(\alpha-2)(\alpha-3)}$$

The expression for skewness is given as

$$\frac{2U^3 - 3U_1U_2 + U_3}{U_2^{\frac{3}{2}}}$$

Where

$$2U_1^3 = 2\left(\frac{\lambda}{(\alpha-1)}\right)^3 = \frac{2\lambda^3}{(\alpha-1)^3}$$

$$3U_1U_2 = 3\left(\frac{\lambda}{(\alpha-1)}\right)\left(\frac{2\lambda^2}{(\alpha-1)(\alpha-2)}\right) = 3\left(\frac{2\lambda^3}{(\alpha-1)^2(\alpha-2)}\right)$$

$$U_3 = \frac{6\lambda^3}{(\alpha-1)(\alpha-2)(\alpha-3)}$$

$$U_2^{\frac{3}{2}} = \sqrt{\left(\frac{\alpha\lambda^2}{(\alpha-1)^2(\alpha-2)}\right)^3} = \frac{\sqrt{(\alpha^3\lambda^6)}}{(\alpha-1)^6(\alpha-2)^3}$$

Inserting all into the expression for skewness

$$\frac{\frac{2\lambda^3}{(\alpha-1)^3} - \frac{6\lambda^3}{(\alpha-1)^2(\alpha-2)} + \frac{6\lambda^3}{(\alpha-1)(\alpha-2)(\alpha-3)}}{\frac{\sqrt{\alpha^3\lambda^6}}{(\alpha-1)^6(\alpha-2)^3}}$$

Simplifying further gives

$$\frac{2(1+\alpha)}{\alpha-3}\sqrt{\frac{\alpha-2}{\alpha}} \qquad \alpha > 3$$

The expression is the third moment or skewness of the pareto distribution

The fourth Moment $r = 4$ (Kurtosis)

We can obtain the Kurtosis which is the peakedness of the distribution by

putting $r = 4$ in equation 5 and simplifying

$$U_4^1 = \alpha\lambda^2\left[\frac{\Gamma(\alpha-4)\Gamma 4+1}{\Gamma\alpha+1}\right]$$

$$\alpha\lambda^4\left[\frac{\Gamma(\alpha-4)\,\Gamma 5}{\Gamma\alpha+1}\right]$$

$$\Gamma 5 = 24$$

$$\Gamma(\alpha-4) = (\alpha-4-1)! = (\alpha-5)!$$

$$\Gamma(\alpha+1) = \alpha!$$

$$U_4^1 = 24\alpha\lambda^4 \left[\frac{(\alpha-5)!}{\alpha!}\right]$$

$$\frac{24\alpha\lambda^4(\alpha-5)}{\alpha(\alpha-1)(\alpha-2)(\alpha-3)(\alpha-4)(\alpha-5)}$$

Dividing like terms

$$\frac{24\lambda^4}{(\alpha-1)(\alpha-2)(\alpha-3)(\alpha-4)}$$

The expression for Kurtosis (Spearsons moment) is given as

$$\frac{U_4 - 4U_1U_3 + 6U_1^2U_2 - 3U_1^4}{Var^2(X)}$$

Where

$$U_4 = \frac{24\lambda^4}{(\alpha-1)(\alpha-2)(\alpha-3)(\alpha-4)}$$

$$4U_1U_3 = 4\left(\frac{\lambda}{(\alpha-1)}\right)\left(\frac{6\lambda^3}{(\alpha-1)(\alpha-2)(\alpha-3)}\right)$$

$$= 4\left(\frac{6\lambda^4}{(\alpha-1)^2(\alpha-2)(\alpha-3)}\right)$$

$$6U_1^2U_2 = 6\left(\frac{\lambda^2}{(\alpha-2)^2}\right)\left(\frac{2\lambda^2}{(\alpha-1)(\alpha-2)}\right)$$

$$= 6\left(\frac{2\lambda^4}{(\alpha-1)^3(\alpha-2)}\right)$$

$$3U_1^4 = 3\left(\frac{\lambda}{(\alpha-1)}\right)^4 = \frac{3\lambda^4}{(\alpha-1)^4}$$

$$Var^2(X) = \left(\frac{\alpha\lambda^2}{(\alpha-1)^2(\alpha-2)}\right)^2$$

Inserting all into the expression for Kurtosis stated above

$$\frac{\frac{24\lambda^4}{(\alpha-1)(\alpha-2)(\alpha-3)(\alpha-4)} - 4\left(\frac{6\lambda^4}{(\alpha-1)^2(\alpha-2)(\alpha-3)}\right) + 6\left(\frac{2\lambda^4}{(\alpha-1)^3(\alpha-2)}\right) - 3\left(\frac{\lambda^4}{(\alpha-2)^4}\right)}{\left(\frac{\alpha\lambda^4}{(\alpha-1)^2(\alpha-2)}\right)^2}$$

From further simplification, we obtain the kurtosis of the distribution

given as

$$U_4^1 = \frac{6(\alpha^3+\alpha^2-6\alpha-2)}{\alpha(\alpha-3)(\alpha-4)} \qquad \alpha > 4$$

3.5. SUMMARY

In this chapter we have been able to find expresions for the raw moment, the first, second, third and fourth moments using the knowledge of calculus and transformation of variables.

CHAPTER FOUR

ANALYSIS AND APPLICATION

4.0. INTRODUCTION

In this chapter, we shall take a look at analysis and aplication of methods of results that were obtained in the preceeding chapters. The aim is to choose numerical values for the parameters that fits into the raw moment.

A value is choosen for a parameter and kept constant at different levels while varying the other parameter value to observe the trend in the across the first, second, third and fourth moments of the distribution under study at the differents levels. Also, we shall find the mean and variance, skewness and kurtosis. This is basically the essence of this chapter and will be presented in a tabular form to give a simplistic, reader friendly understanding of this chapter.

4.1. ANALYSIS AND INTERPRETATION

$\alpha > r \; \forall \; \alpha \neq r$

r=1-4, α= 4.5-9, λ = 1.5 (constant)

Table 1.0	**Parameter Values**					
Moments	α=4.5, λ=1.5	α=5, λ=1.5	α=6, λ=1.5	α=7, λ=1.5	α=8, λ=1.5	α=9, λ=1.5
U_1	0.4286	0.3750	0.3000	0.2500	0.2143	0.1875
U_2	0.5142	0.3750	0.2250	0.1500	0.1071	0.0803
U_3	1.5428	0.8438	0.3375	0.1686	0.0964	0.0603
U_4	18.5143	5.0625	1.0125	0.3375	0.1446	0.0723

Table 1.1

Mean	0.4286	0.3750	0.3000	0.2500	0.2143	0.1875
Variance	0.3306	0.2344	0.1350	0.0875	0.0612	0.0452
Skewness	5.4689	4.6480	3.8103	3.3748	3.1213	2.397
Kurtosis	149.4444	73.8000	38.6667	27.8571	22.7250	16.7556

From the tables above, the first table represents the outcomes of the individual moments when the parameters are assigned values. In this case, λ is held constant at 1.5 while α varies from 4.5, 5, 6, 7, 8 and 9.

Taking a very close look at the table, we observe a trend or pattern in the outcome. At the first colum where α =4.5 and λ = 1.5 we can see an increase in value from the first moment down the fourth. At the second column, we observe that the first moment and the second takes the same values, shortly followed by a trmendous increase from the 3rd moment to the 4th.

At the 3rd and 4th columns, we observe a similar trend between both, such that there is a decrease from the 1st to 2nd moment and a sudden increase from the 3rd down to the 4th moment. The same pattern is observed in the 5th and last columns but this time a decrease from the 1st to 3rd moment and an increase at the 4th moment.

Meanwhile, at the rows for each moments, there is a decrease for α=4.5 to 9 and λ= 1.5. The second table provides values for the mean, variance, skewnes and kurtosis for each column of the first table. From observation, a similar trend follows for each column of the table. A decline between the mean and variance followed by an lncrease from the skewness to the kurtosis.

Table for r=1,2,3,4 α= 10,11,13-16, λ = 3 (constant) $\alpha > r \, \forall \, \alpha \neq r$

Table 1.2	**Parameter Values**					
Moments	α=10, λ=3	α=11, λ=3	α=13, λ=3	α=14, λ=3	α=15, λ=3	α=16, λ=3
U_1	0.3333	0.3000	0.2500	0.2308	0.2143	0.2000
U_2	0.2500	0.2000	0.1363	0.1154	0.0989	0.0857
U_3	0.3214	0.2250	0.1227	0.0944	0.0742	0.0593
U_4	0.6428	0.3857	0.1636	0.1133	0.0809	0.0593

Table 1.3

Mean	0.3333	0.3000	0.2500	0.2308	0.2143	0.2000
Variance	0.1389	0.1100	0.0739	0.0621	0.0529	0.0457
Skewness	2.8111	2.7136	2.5756	2.5249	2.4825	2.4465
Kurtosis	17.8286	16.4805	14.7231	14.1195	13.6303	13.2259

r=1,2,3,4 α= 17-22, λ = 4 (constant) $\alpha > r \; \forall \; \alpha \neq r$

Table 1.4

Moments	Parameter Values					
	α=17, λ=4	α=18, λ=4	α=19, λ=4	α=20, λ=4	α=21, λ=4	α=22, λ=4
U_1	0.2500	0.2353	0.2222	0.2105	0.2000	0.1904
U_2	0.1333	0.1176	0.1046	0.0936	0.0842	0.0761
U_3	0.1142	0.0941	0.0784	0.0660	0.0561	0.0481
U_4	0.1407	0.1076	0.0837	0.0660	0.0528	0.0428

Table 1.5

Mean	0.2500	0.2353	0.2222	0.2105	0.2000	0.1904
Variance	0.0708	0.0622	0.0552	0.0492	0.0442	0.0399
Skewness	2.4154	2.3884	2.3648	2.3438	2.3251	2.3084
Kurtosis	12.8862	12.5968	12.3474	12.1301	11.9393	11.7703

Table 1.2 shows values for increased value of α and λ i.e at α=10, 11, 13, 14, 15, 16 and λ=3 followed by a table for it mean, variance and skewness. From the table, we observe trends in the rows and columns similar to the ones earlier obtained. At the first column, there is a decrease between the first and second moments followed by an increase from the third to the fourth. The same trend is seen in the second column, but at the third colum, we see a decrease from the first to the third moments.

Meanwhile, across every row there is a decrease in as the valiue of α increases.

Table 1.3 also show values for increased λ at 4 and increasing α from 17 to 22 with it corresponding mean, variance, skewness and kurtosis. The trend follows a similar pattern like table 1.2 i.e decrease in value from the first moment shortly followed by an increase at certain point.

In general, from all the table of α and λ we can make the following assumptions:

- If the values of α is relatively small, like in the first table, the values of each moment at each columns are large but as α gets larger (comparing with the succesive tables) the values for the first to the fourth moments declines and tends to Zero.
- At every constant value of λ there is always a decrease across the rows
- If the values of α increases with constant or varying value for λ, the values from first to fourth moment decreases, tending to Zero at the rows and columns.

In previous tables, we try to look at trends that occurred at different values of α while holding λ constant. In subsequent table, we look at a reverse case for a fixed value of α while varying λ and observe the patterns that occur across different tables if the values are increasing, Table for α= 4.5 (constant) λ = 2, 2.5, 3, 3.5, 4, 5 $\alpha > r \; \forall \; \alpha \neq r$

Table 1.6	**Parameter Values**					
Moments	α=4.5 λ=2	α=4.5 λ=2.5	α=4.5 λ=3	α=4.5 λ=3.5	α=4.5 λ=4	α=4.5 λ=5
U_1	0.5714	0.7143	0.8571	1.0000	1.1428	1.4287
U_2	0.9143	1.4286	2.0571	2.8000	3.6571	5.7143
U_3	3.6571	7.1428	12.3428	19.6000	29.2571	57.1428
U_4	58.5143	142.8571	296.2286	548.8000	936.2286	2285.7140

Table 1.7

Mean	0.5714	0.7143	0.8571	1.0000	1.1428	1.4287
Variance	0.5878	0.9184	1.3225	1.8000	2.3511	3.6731
Skewness	5.4659	5.4659	5.4659	5.4659	5.4659	5.4659
Kurtosis	149.4444	149.4444	149.4444	149.4444	149.4444	149.4444

Table for r=1,2,3,4 α= 5.5 (constant), λ = 7-12 $\alpha > r \; \forall \; \alpha \neq r$

Table 1.8	**Parameter Values**					
Moments	α=5.5 λ=7	α=5.5 λ=8	α=5.5 λ=9	α=5.5 λ=10	α=5.5 λ=11	α=5.5 λ=12
U_1	1.5556	1.7778	2.0000	2.2222	2.4444	2.6667
U_2	6.2222	8.1269	10.2857	12.6984	15.3651	18.2857
U_3	52.2607	78.0191	111.0857	152.3810	202.8190	263.3143
U_4	975.6444	1664.4060	2666.0570	4063.4920	5949.3590	8426.0507

Table 1.9

Mean	1.5556	1.7778	2.0000	2.2222	2.4444	2.6667
Variance	3.8023	4.9663	6.2857	7.7602	9.3900	11.1744
Skewness	4.1482	4.1482	4.1482	4.1482	4.1482	4.1482
Kurtosis	50.0182	50.0182	50.0182	50.0182	50.0182	50.0182

The tables represents the results of Analysis obtained for assigning values to the parameters α and λ. Unlike the previous tables where λ was held constant while varying α, here the reverse becomes the case, α will be held at constant.

Table 1.6 represents values at $\alpha = 4.5$ and $\lambda = 2, 2.5, 3, 3.5, 4, 5$ from the table, we observe an increase, increasing in row and column as the value of λ increases. We also observe that the values for 3rd and 4th moment are very large compared to the previous tables where λ is constant.

Table 1.7 are values for the mean, variance, skewness and kurtosis, from the table, the trend in value of mean and variance follows the same trend as it table of moments except that the skewness and kurtosis have a constant values throughout, this is due to the fact that the skewness and kurtosis of the Pareto II distribution are void of the paramenter λ

We also examine the trend when the values of α and λ are increased in table 1.8 and 1.9. We see that the values are largely increasing across the rows and columns.

Table 2.0	Parameter Values					
Moments	α=4.5 λ=4.5	α=5, λ=5	α=6 λ=6	α=7 λ=7	α=8 λ=8	α=9 λ=9
U_1	1.2857	1.2500	1.2000	1.1667	1.1428	1.1250
U_2	4.6286	4.1667	3.6000	3.2667	3.0476	2.8929
U_3	41.6571	31.2500	21.6000	17.1500	14.6286	13.0179
U_4	1499.6570	625.0000	259.2000	160.0667	117.0286	93.7286

Table 2.1

Mean	1.2857	1.2500	1.2000	1.1667	1.1428	1.1250
Variance	2.9755	2.6042	2.1600	1.9056	1.7415	1.6272
Skewness	5.4659	4.6476	3.8103	3.3806	3.1177	2.9397
Kurtosis	149.4444	73.8000	38.6667	27.8571	22.7250	19.7556

Table 2.2	Parameter Values					
Moments	α=10 λ=10	α=11 λ=11	α=12 λ=12	α=13 λ=13	α=14 λ=14	α=15 λ=15
U_1	1.1111	1.1000	1.0909	1.0833	1.0769	1.0714
U_2	2.7778	2.6889	2.6182	2.5606	2.5128	2.4725
U_3	11.9048	11.0917	10.4727	9.9864	9.5944	9.2719
U_4	79.3651	69.7191	62.8364	57.6989	53.7287	50.5744

Table 2.3

Mean	1.1111	1.1000	1.0909	1.0833	1.0769	1.0714
Variance	1.5432	1.4789	1.4281	1.3869	1.3531	1.3246
Skewness	2.8111	2.7136	2.6372	2.5756	2.5249	2.4825
Kurtosis	17.8286	16.4805	15.4861	14.7231	14.1195	13.6303

In preceeding tables, we have looked at what happens when a parameter is constant and the other is varied. In the following tables, we want to observe patterns when both parameters α and λ are held constant. Table 2.0 to 2.3 are the values obtained for α and λ from 4.5 to 15

From the table we can see that at each column where α and λ are constant, the values from the first to fourth moment along the column is increasing meanwhile across the row for each individual moments the values are decreasing. The same trend is observed in the tables that contain the mean, variance, skewness and kurtosis. This is quite similar to table 1.0 where α was varied and λ was constant

4.2. DISCUSSION ON FINDINGS

Having applied values to study the behaviour of moments partaining to the Pareto II distribution, we hereby make a discussionary view of our findings. Based on information obtained from table 1.0 as the value of α increases across the row, the values for individual moments bgins to decrease, if this continues, as

$\alpha \to \infty$, the 4th moment approaches zero. In table 1.6, if λ is increasing, the values are increasing both row and columnwise

Generally, we can make the following few assumptions;

- At any constant value of λ, there is always a decrease across the row
- We can say that α is responsible for the decrease in value across the row for each moment if λ is constant.
- The skewness and kurtosis are not greatly affected when λ is varied at a constant α.
- The increase in value across the row for each moment can be attributed to the parameter λ.

4.3. SUMMARY

In this chapter, we have been able to look at the behaviour of moments partaining to the Pareto II distribution by applying values to parameters and observing the trend that occurs with each moments and with the mean, variance, skewness and kurtosis. With the help of Statiscal Software "R"

CHAPTER FIVE

SUMMARY AND CONCLUSION

5.0. SUMMARY

This chapter concludes the work with a summary of work done from the preceeding chapters. So far, we have focused on the study of moment of Pareto II distribution. The nascence of this work was set out in chapter one with an introduction to moments, the aim and objectives to be accomplished, significance of the study and definition of some basic terms.

In the next chapter, we were able to review some literature done by scholars generally on Pareto Distribution, we also looked at the methodology in chapter three where we obtained various expressions using the knowledge of differential and integral calculus combined with transformation of variables, we were able to obtain the probability density function given it Cummulative density function, we then moved on to obtain the expression for the raw moment using the knowledge of integration and transformation of variable from the beta function which enabled us to obtain expressions for the first, second, third and fourth moment when r=1,2,3,4 respectively.

In chapter four, we dealt with analysis and application of methods obtained in chapter three. We obsereved the behaviour of each moments by choosing numerical values for the parameters of the distribution and also discuss the findings.

5.1. CONCLUSION

One of the most important aspect of probability distribution is the study of moments. It is very important and applicable to probability distributions as we have seen in this study. We believe that a thorough study of this can assist in various ways such as describing the measure of centre of data as in the mean, it could also help in describing the variabiility of the data from the centre of location.

Study of moment is also important in determining the normality of distribution by information obtained from the skewness and kurtosis. A normal distribution will have a skewness of zero, if the kurtosis is greater than 3 then the dataset has heavier tails than a normal distribution.

In this regard, Pareto distribution, originally applied to describing the distribution of wealth in society has found it application in description of social,quality control, scientific, geophysical, actuarial and many other phenomena.

REFERENCES

Abdell-All et al (2003) Geometric Properties of Pareto Distribution. *Applied Mathematics and computation, volume 145, Issue 2-3, Pg 321-339*

Alasdair, G.M (2007) Vilfredo Pareto's Sociology; *A framework for political psychology (Rethinking classical sociology) Glasgow Caledonian University, UK*

Arnold, B.C (2014) Pareto distribution, *Riverside California*

Atkinson A, Harrison A (1978). Distribution of personal wealth in Britain. Cambridge University Press, Cambridge

Chahkandi, M. and Ganjali, M. (2009). on some lifetime distributions with decreasing failure rate, *Computational Statistics and Data Analysis*, 53: 4433–4440.

Champernowne, D.G (1937) The Theory of Income Distribution. *Econometrica*

Chu, J et al (2019) A review of goodness of fit test for Pareto distributions, *journal of Computational and Applied Mathematics*

Corbellini A, Crosato L, Ganugi P, Mazzoli M (2007). Fitting Pareto II distributions on firm size: Statistical methodology and economic puzzles. Paper presented at the international conference on applied stochastic models and data analysis, Chania, Crete.

Cordeiro G, Ortega E, Popović B (2013) The gamma-Lomax distribution. J Stat Comput Simul 85(2):305–319

Cordeiro, G. M., Ortega, E. M. M., Popovic, B. V and Pescim, R. R. (2014). The Lomax generator of distributions: Properties, minification process and regression model, *Applied Mathematics and Computation,* 247:465-486.

Francois Geerolf (2016) A Theory of Pareto distributions

Marcelo, B et al The kumaraswary Pareto Distribution, *Departmento de estatistia, Universidade Federal de Pernambuco*

Osagie S.A. and Nosakhare F.U. (2020). Properties and application of a new three-parameter Lomax-Exponential distribution. Journal of the Nigerian Association of Mathematical Physics

Pareto, Vilfredo (1898). "Cours d'economie politique". *Journal of Political Economy*. **6**. doi:10.1086/250536.

Punathumparambath B (2011). Estimation of P(X > Y) for the double Lomax distribution. Prob Stat Forum 4:1–11

Sinch, V.P & Guo, H (1995) Parameter estimation for 3-parameter generalized pareto distribution by the principle of maximum entropy

Tadikamalla, P. R. (1980). A look at the Burr and related distributions, *International Statistics Review,* 48: 337–344.

APPENDIX

```
> m1<-rmoments(1,4.5,1.5)
> m1
[1] 0.4285714
> m2
[1] 0.5142857
> m3
[1] 1.542857
> m4
[1] 18.51429
> variance
[1] 0.3306122
> skewness
[1] 5.465944
> kurtosis
[1] 149.4444

> m1<-rmoments(1,5,1.5)
> m2<-rmoments(2,5,1.5)
> m3<-rmoments(3,5,1.5)
> m4<-rmoments(4,5,1.5)
> m1
[1] 0.375
> m2
[1] 0.375
> m3
[1] 0.84375
> m4
[1] 5.0625
> variance
[1] 0.234375
> skewness
[1] 4.64758
> kurtosis
[1] 73.8
> m1<-rmoments(1,6,1.5)
> m2<-rmoments(2,6,1.5)
> m3<-rmoments(3,6,1.5)
> m4<-rmoments(4,6,1.5)
> m1
[1] 0.3
> m2
[1] 0.225
> m3
[1] 0.3375
> m4
[1] 1.0125
> variance
[1] 0.135
> skewness
[1] 3.810317
> kurtosis
[1] 38.66667
> m1<-rmoments(1,7,1.5)
> m2<-rmoments(2,7,1.5)
> m3<-rmoments(3,7,1.5)
> m4<-rmoments(4,7,1.5)
> m1
[1] 0.25
> m2
[1] 0.15
> m3
[1] 0.16875
> m4
[1] 0.3375
> variance
[1] 0.0875
> skewness
[1] 3.380617
> kurtosis
[1] 27.85714

> m1<-rmoments(1,8,1.5)
> m2<-rmoments(2,8,1.5)
> m3<-rmoments(3,8,1.5)
> m4<-rmoments(4,8,1.5)
> m1
[1] 0.2142857
> m2
[1] 0.1071429
> m3
[1] 0.09642857
> m4
[1] 0.1446429
> variance
[1] 0.06122449
> skewness
[1] 3.117691
> kurtosis
[1] 22.725
> m1<-rmoments(1,9,1.5)
> m2<-rmoments(2,9,1.5)
> m3<-rmoments(3,9,1.5)
> m4<-rmoments(4,9,1.5)
> m1
[1] 0.1875
> m2
[1] 0.08035714
> m3
[1] 0.06026786
> m4
[1] 0.07232143
> variance
[1] 0.04520089
> skewness
[1] 2.939724
> kurtosis
[1] 19.75556

> m1<-rmoments(1,10,3)
> m2<-rmoments(2,10,3)
> m3<-rmoments(3,10,3)
> m4<-rmoments(4,10,3)
> m1
[1] 0.3333333
> m2
[1] 0.25
> m3
[1] 0.3214286
> m4
[1] 0.6428571
> variance
[1] 0.1388889
> skewness
[1] 2.811057
> kurtosis
[1] 17.82857

> m1<-rmoments(1,11,3)
> m2<-rmoments(2,11,3)
> m3<-rmoments(3,11,3)
> m4<-rmoments(4,11,3)
> m1
[1] 0.3
> m2
[1] 0.2
> m3
[1] 0.225
> m4
[1] 0.3857143
> variance
[1] 0.11
> skewness
[1] 2.713602
> kurtosis
[1] 16.48052

> m1<-rmoments(1,13,3)
> m2<-rmoments(2,13,3)
> m3<-rmoments(3,13,3)
> m4<-rmoments(4,13,3)
> m1
[1] 0.25
> m2
[1] 0.1363636
> m3
[1] 0.1227273
> m4
[1] 0.1636364
> variance
[1] 0.07386364
> skewness
[1] 2.575625
> kurtosis
[1] 14.72308
>
```

```
> m1<-rmoments(1,14,3)
> m2<-rmoments(2,14,3)
> m3<-rmoments(3,14,3)
> m4<-rmoments(4,14,3)
> m1
[1] 0.2307692
> m2
[1] 0.1153846
> m3
[1] 0.09440559
> m4
[1] 0.1132867
> variance
[1] 0.06213018
> skewness
[1] 2.524964
> kurtosis
[1] 14.11948

> m1<-rmoments(1,15,3)
> m2<-rmoments(2,15,3)
> m3<-rmoments(3,15,3)
> m4<-rmoments(4,15,3)
> m1
[1] 0.2142857
> m2
[1] 0.0989011
> m3
[1] 0.07417582
> m4
[1] 0.08091908
> variance
[1] 0.05298273
> skewness
[1] 2.482532
> kurtosis
[1] 13.6303

> m1<-rmoments(1,16,3)
> m2<-rmoments(2,16,3)
> m3<-rmoments(3,16,3)
> m4<-rmoments(4,16,3)
> m1
[1] 0.2
> m2
[1] 0.08571429
> m3
[1] 0.05934066
> m4
[1] 0.05934066
> variance
[1] 0.04571429
> skewness
[1] 2.446468
> kurtosis
[1] 13.22596

> m1<-rmoments(1,17,4)
> m2<-rmoments(2,17,4)
> m3<-rmoments(3,17,4)
> m4<-rmoments(4,17,4)
> m1
[1] 0.25
> m2
[1] 0.1333333
> m3
[1] 0.1142857
> m4
[1] 0.1406593
> variance
[1] 0.07083333
> skewness
[1] 2.415437
> kurtosis
[1] 12.88623
> m1<-rmoments(1,18,4)
> m2<-rmoments(2,18,4)
> m3<-rmoments(3,18,4)
> m4<-rmoments(4,18,4)
> m1
[1] 0.2352941
> m2
[1] 0.1176471
> m3
[1] 0.09411765
> m4
[1] 0.107563
> variance
[1] 0.06228374
> skewness
[1] 2.38845
> kurtosis
[1] 12.59683
> m1<-rmoments(1,19,4)
> m2<-rmoments(2,19,4)
> m3<-rmoments(3,19,4)
> m4<-rmoments(4,19,4)
> m1
[1] 0.2222222
> m2
[1] 0.1045752
> m3
[1] 0.07843137
> m4
[1] 0.08366013
> variance
[1] 0.05519245
> skewness
[1] 2.364763
> kurtosis
[1] 12.34737
> m1<-rmoments(1,20,4)
> m2<-rmoments(2,20,4)
> m3<-rmoments(3,20,4)
> m4<-rmoments(4,20,4)
> m1
[1] 0.2105263
> m2
[1] 0.09356725
> m3
[1] 0.06604747
> m4
[1] 0.06604747
> variance
[1] 0.04924592
> skewness
[1] 2.343806
> kurtosis
[1] 12.13015
> m1<-rmoments(1,21,4)
> m2<-rmoments(2,21,4)
> m3<-rmoments(3,21,4)
> m4<-rmoments(4,21,4)
> m1
[1] 0.2
> m2
[1] 0.08421053
> m3
[1] 0.05614035
> m4
[1] 0.05283798
> variance
[1] 0.04421053
> skewness
[1] 2.32513
> kurtosis
[1] 11.93931
> m1<-rmoments(1,22,4)
> m2<-rmoments(2,22,4)
> m3<-rmoments(3,22,4)
> m4<-rmoments(4,22,4)
> m1
[1] 0.1904762
> m2
[1] 0.07619048
> m3
[1] 0.0481203
> m4
[1] 0.0427736
> variance
[1] 0.0399093
> skewness
[1] 2.308383
> kurtosis
[1] 11.77033
```

```
> m1<-rmoments(1,4.5,2)
> m2<-rmoments(2,4.5,2)
> m3<-rmoments(3,4.5,2)
> m4<-rmoments(4,4.5,2)
> m1
[1] 0.5714286
> m2
[1] 0.9142857
> m3
[1] 3.657143
> m4
[1] 58.51429
> variance
[1] 0.5877551
> skewness
[1] 5.465944
> kurtosis
[1] 149.4444
> m1<-rmoments(1,4.5,2.5)
> m2<-rmoments(2,4.5,2.5)
> m3<-rmoments(3,4.5,2.5)
> m4<-rmoments(4,4.5,2.5)
> m1
[1] 0.7142857
> m2
[1] 1.428571
> m3
[1] 7.142857
> m4
[1] 142.8571
> variance
[1] 0.9183673
> skewness
[1] 5.465944
> kurtosis
[1] 149.4444

> m1<-rmoments(1,4.5,3)
> m2<-rmoments(2,4.5,3)
> m3<-rmoments(3,4.5,3)
> m4<-rmoments(4,4.5,3)
> m1
[1] 0.8571429
> m2
[1] 2.057143
> m3
[1] 12.34286
> m4
[1] 296.2286
> variance
[1] 1.322449
> skewness
[1] 5.465944
> kurtosis
[1] 149.4444
>
>
> m1<-rmoments(1,4.5,3.5)
> m2<-rmoments(2,4.5,3.5)
> m3<-rmoments(3,4.5,3.5)
> m4<-rmoments(4,4.5,3.5)
> m1
[1] 1
> m2
[1] 2.8
> m3
[1] 19.6
> m4
[1] 548.8
> variance<-(m2-m1^2)
> skewness<-(((2*m1^3)-(3
*m1*m2)+m3)/((m2-m1^2)^(3
/2)))
> kurtosis<-((m4-(4*m1*m3
)+(6*m1^2*m2)-(3*m1^4))/(
m2-m1^2)^2)
> variance
[1] 1.8
> skewness
[1] 5.465944
> kurtosis
[1] 149.4444
> m1<-rmoments(1,4.5,4)
> m2<-rmoments(2,4.5,4)
> m3<-rmoments(3,4.5,4)
> m4<-rmoments(4,4.5,4)
> m1
[1] 1.142857
> m2
[1] 3.657143
> m3
[1] 29.25714
> m4
[1] 936.2286
> variance
[1] 2.35102
> skewness
[1] 5.465944
> kurtosis
[1] 149.4444
> m1<-rmoments(1,4.5,5)
> m2<-rmoments(2,4.5,5)
> m3<-rmoments(3,4.5,5)
> m4<-rmoments(4,4.5,5)
> m1
[1] 1.428571
> m2
[1] 5.714286
> m3
[1] 57.14286
> m4
[1] 2285.714
> variance
[1] 3.673469
> skewness
[1] 5.465944
> kurtosis
[1] 149.4444
```